Robust Derivation of Risk Reduction Strategies[1][2]

Julian Richardson, RIACS/USRA, NASA Ames Research Center, julianr@email.arc.nasa.gov
Daniel Port, University of Hawaii, Dept. of Information Technology Management, dport@hawaii.edu
Martin Feather, Jet Propulsion Laboratory, Caltech, martin.s.feather@jpl.nasa.gov

Abstract—Effective risk reduction strategies can be derived mechanically given sufficient characterization of the risks present in the system and the effectiveness of available risk reduction techniques. Quantitative assessments of risks and risk reduction techniques are likely to be inaccurate. In this paper we describe sensitivity analysis experiments which we carried out to evaluate how inaccurate quantification of risk and risk reduction techniques affect the performance of mechanically derived risk reduction strategies. Our experiments show that mechanically derived risk reduction strategies are likely to produce significant improvements in risk reduction compared to a alternative risk reduction strategies, and arguably should be used as a matter of course, even when knowledge of risk and risk reduction techniques is very inaccurate.

TABLE OF CONTENTS

1. INTRODUCTION

Effective risk reduction strategies can be derived mechanically given sufficient characterization of the risks present in the system and the effectiveness of available risk reduction techniques.

The Strategic Method [Port et al, 2005] is one technique for generating optimal risk reduction strategies, which most reduce risk exposure for a given budget. The input to the strategic method is a specification, for the risks and risk reduction techniques of interest, of the probability and cost of failure for each risk, both before and after application of each risk reduction technique, and the cost of applying each risk reduction technique.

In this paper, we address an important question: can we reliably expect mechanically derived risk reduction strategies to be better than fixed or hand-selected risk reduction strategies, given that the quantitative assessment of risks and risk reduction techniques upon which mechanical derivation is based is likely to be inaccurate?

We consider this question relative to two methods for deriving effective risk reduction strategies: the Strategic Method and the Defect Detection and Prevention (DDP) tool [Feather & Cornford, 2003]. We evaluate the efficacy of the Strategic Method as follows:

1. How much more does the strategy computed by the Strategic Method reduce risk than a fixed strategy?

2. How does the accuracy with which we can assess the risk reduction achievable by each technique affect the performance of the strategic method?

3. How effectively does the strategic method tailor its recommendations to the specific levels of risk present in a project?

We find that:

1. The strategic method performs significantly better than a reasonable strategy which adopts techniques in order of cheapness, and than random strategies.

2. The strategic method is effective even when there is large uncertainty on the risk reduction achievable by the available risk reduction methods.

3. Strategies computed by the strategic method are effectively tailored to projects' specific risk levels.

4. Strategies computed by the Strategic Method are at least as good as those computed by a simulated annealing optimization method implemented in DDP.

[1] 1-4244-0525-4/07/$20.00 ©2007 IEEE.
[2] Paper 1346 Revision number 6.

2. THE STRATEGIC METHOD

Any development activity and system operation involves risk. Risks are possible situations that can cause a system to fail to meet its goals. They range in impact from trivial to fatal and in likelihood from certain to improbable. Generally risks are either "identified" in that they arise from anticipated system errors and off-nominal conditions or "unidentified" where they do not. Furthermore the impact of identified risks are either "known" where the expected loss-potential has been assessed or "unknown" where the loss-potential has not or cannot be assessed. Risks that are unidentified or have unknown impacts are sometimes loosely labeled as "risks due to uncertainty". Risk considerations often focus on uncertainty since typically known risks are either addressed or accepted as within a "tolerable" level. A *risk model* describes risks and their impacts for a particular system.

Risks are typically not static. Likelihoods and impacts change with a number of dynamic variables, e.g. time, cost, system state. As a consequence it is often desirable to consider risks with respect to a planned set of events such as assessment effort, system operation time, development investment, etc. We call the representation of risks that change dynamically over planned activities a *strategic risk model*. We use the term "strategic" here because there is an implicit planned order of risk reduction activities – a *strategy* – that is expected to achieve a specified goal.

We use as a basic measure of risk the *risk exposure* (RE), which is computed as the product of the probability of loss P(L) and size that loss S(L) summed over all sources of loss for a particular risk item. Total system risk exposure is the sum of individual risk exposures, total RE = $\sum P(L_i)* S(L_i)$, where L_i is the loss due to the i^{th} risk.

Related to the notion of RE is risk reduction leverage (RRL). RRL is a way of gauging the effectiveness or desirability of a risk reduction technique. If RRCost stands for the cost of the activity that achieves the risk reduction, then the formula for RRL is:

$$RRL = (RE_{before} - RE_{after}) / RRCost$$

It is often the case that a technique reduces only the likelihood of a risk and not its magnitude. In this case, the RRL reduces to the cost-benefit (CB) ratio:

$$CB = [P_{before}(L) - P_{after}(L)]*S(L)/RRCost$$
$$= \Delta P(L)*S(L)/RRCost$$

Since risk considerations are critical to the success or failure of a system, it is important that risks be investigated candidly and completely. A *risk profile* (or RE profile) is the evaluation of RE as a function of a monotonically increasing quantity such as elapsed time, cumulative effort, or cumulative cost.

For example, Table 1 shows typical sets of risk and cost data provided for calculating the effectiveness of IV&V assessment techniques (see Appendix for descriptions of attributes and techniques used here):

Attribute (i)	A1	A2	A3	A4	A5	A6	A7
Loss potential for Ai	100	90	90	80	60	30	50
P_{before}(Ai)	6	5	20	15	20	5	20
Attribute (i)	A8	A9	A10	A11	A12	A13	A14
Loss potential for Ai	20	10	10	60	10	90	60
P_{before}(Ai)	10	10	10	30	20	50	40

Table 1: Attributes and their Loss Potential and Probability (before mitigation)

Cost to assess A_i w/ T_j	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
T1	50	x	10	70	10	x	x	x	x	50	5	x	10	x
T2	100	x	x	100	100	x	x	x	x	x	x	x	x	x
T3	x	x	80	80	80	x	x	x	x	x	x	x	x	x
T4	100	90	x	x	19	x	x	x	x	x	x	x	x	x
T5	70	100	70	70	70	x	x	x	x	x	x	x	x	x
T6	30	30	30	30	30	x	x	x	x	x	x	x	x	x
T7	x	x	x	x	x	5	10	x	5	5	3	x	3	x
T8	x	x	x	x	x	80	70	x	80	80	x	x	x	x
T9	x	x	x	x	x	x	3	10	20	20	20	10	20	10
T10	60	x	x	60	50	40	50	50	50	40	40	20	40	20
T11	60	x	90	60	60	x	x	x	x	50	10	x	10	x
T12	x	x	x	x	x	5	5	10	10	10	10	5	x	x
T13	30	x	x	30	30	x	x	30	x	30	5	x	30	x
T14	100	x	x	100	100	x	x	x	x	100	5	x	100	x

Table 2: Risk Assessment Techniques and the Costs of Assessing Them

P_{after}(Ai) using Tj	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14
T1	4	x	15	12	15	x	x	x	x	5	15	x	20	x
T2	6	x	x	13	15	x	x	x	x	x	x	x	x	x
T3	x	x	15	12	13	x	x	x	x	x	x	x	x	x
T4	6	0	x	x	19	x	x	x	x	x	x	x	x	x
T5	6	2	2	13	18	x	x	x	x	x	x	x	x	x
T6	6	2	5	13	19	x	x	x	x	x	x	x	x	x
T7	x	x	x	x	x	2	15	x	8	10	30	x	30	x
T8	x	x	x	x	x	1	10	x	7	9	x	x	x	x
T9	x	x	x	x	x	x	10	4	6	8	25	20	30	30
T10	6	x	x	12	19	3	15	8	8	8	27	20	30	20
T11	3	x	15	5	5	x	x	x	x	5	5	x	5	x
T12	x	x	x	x	x	3	18	9	10	10	30	20	x	x
T13	5	x	x	12	15	x	x	5	x	6	20	x	28	x
T14	3	x	x	3	5	x	x	x	x	5	10	x	20	x

"x" : technique can't be used for attribute

Table 3: Probability of a Loss *After* Assessing with Technique Tj

The data in Table 3 was generated by considering each system attribute A_i in Table 2 for $S(A_i)$ in terms of the percentage of the project value that would result from an error in the system attribute and $P_{before}(A_i)$ the corresponding probability (as a percentage) of such a loss occurring. Then if the attribute A_i is assessed with technique T_j the resulting $P_{after}(A_i)$ and the corresponding cost for performing this assessment. The cost is effort in hours used to perform the assessment of the attribute. A strategy is an ordered sequence of attribute-technique pairs $< A_i , T_j >$. Given this information, we can perform the following algorithm to calculate an optimal risk reduction strategy:

Step 1: Identify the most significant system assessment attributes. Label them $A_1 ,\ldots, A_n$

Step 2: Identify the most significant assessment techniques (e.g. product testimonials, prototyping, etc.) applicable to the project, available resources (e.g. staff skills, tools). Label them $T_1 ,\ldots, T_n$

Step 3: Estimate the relative probabilities $P(A_i)$ and size $S(A_i)$ quantities for potential losses associated with attributes i=1,…,n before any assessment. $RE(A_i)$ may be estimated directly, e.g. from historical data.

Step 4: Estimate the cost $C(A_i,T_j)$, size $S(A_i,T_j)$, and probability $P(A_i,T_j)$ after assessment of A_i with T_j and the change in risk exposures $\Delta RE(A_i,T_j) = P(A_i)*S(A_i) - S(A_i,T_j)* P(A_i,T_j)$

Step 5: Calculate the benefit matrix $B(A_i,T_j) = \Delta RE(A_i,T_j) - C(A_i,T_j)$. For each A_i find the T_k where $B(A_i,T_k)$ is maximum. Set $c_i = k$.

Step 6: Using c_i as above, calculate the RRL list, $RRL(A_i,Tc_i) = \Delta RE(A_i,Tc_i) / C(A_i,Tc_i)$ for i =1,…,n. Let V(k) be the index where $RRL(A_{V(k)},Tc_{V(k)})$ is the k^{th} largest element in the RRL list. For example, V(1) corresponds to where $RRL(A_i ,Tc_i)$ is maximum over all i, and T(n) is the minimum.

Step 7: Graph the cumulative RE drop, $RE(n) = RE_{total} - \sum\Delta RE(A_{V(k)}, Tc_{V(k)})$ versus cumulative effort $C(n) = \sum C(A_{V(k)}, Tc_{V(k)})$.

The strategy dictates that one should perform $<A_{V(k)}, Tc_{V(k)}>$ for k =1,2,3,… in this order until the cost outweighs the benefit (i.e. $RRL(A_{V(k)}, Tc_{V(k)}) < 1$) unless other specific risk reduction goals are desired. The strategy generated is one that satisfies the utility function:

$$\min_{\tau,J}\left[\sum_{i=1}^{k} RE_{after}(A_i, T_{J(i)}) + RRCost(A_i, T_{J(i)}) + \sum_{i=K+1}^{N} RE_{before}(A_i)\right]$$

where the minimum is taken over the sets $\{(A_1, T_{J(1)}, \tau(1)), (A_2, T_{J(2)}, \tau(2)), \ldots, (A_N, T_{J(N)}, \tau(3))\}$ and all permutations τ of {1,2,…,N} and functions $J:\{1,2,\ldots,N\}\rightarrow\{1,2,\ldots,M\}$ (i.e. J is a set of non-distinct integers 1 through N). The utility function chooses the attribute-technique pairs that minimizes total RE and cost after k activities have been performed assuming k is arbitrary (i.e. the budget or schedule is unknown or may be cut at any time).

Table 4 (reproduced from [Port et al, 2005]) shows the optimal strategy computed for the matrices in Tables 1-3.

Attrib	Tech	RE Change	Cost	Benefit	CB	risk reduction	cumulative cost
None	None	None	None	None	None	15700	0
A13	T11	4050	10	4040	405	11650	10
A7	T9	500	3	497	167	11150	13
A11	T11	1500	10	1490	150	9650	23
A14	T10	1200	20	1180	60	8450	43
A3	T5	1620	70	1550	23.1	6830	113
A6	T7	90	5	85	18	6740	118
A5	T11	900	60	840	15	5840	178
A8	T9	120	10	110	12	5720	188
A4	T14	960	100	860	9.6	4760	288
A2	T4	450	90	360	5	4310	378
A1	T11	300	60	240	5	4010	438
A9	T9	40	20	20	2	3970	458
A10	T13	40	30	10	1.33	3930	488

Table 4: The optimal strategy computed by the strategic method for the matrices from Tables 1-3.

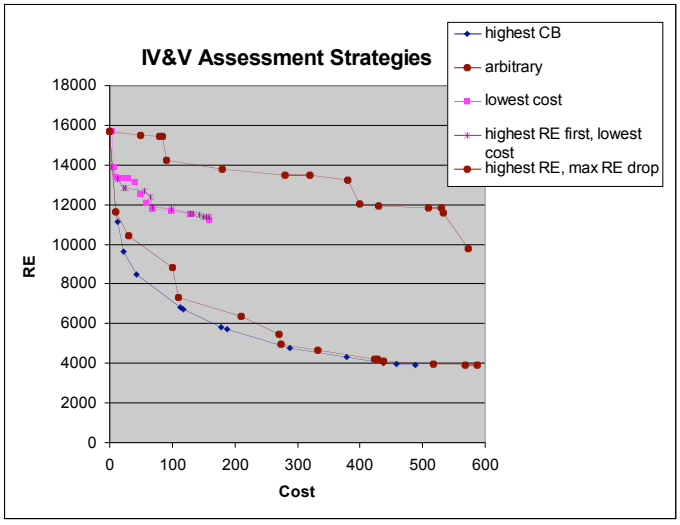

Figure 1: Performance of strategic method strategy versus baseline strategies.

3. STRATEGIC METHOD ROBUSTNESS

In this section, we carry out a systematic evaluation of strategies recommended by the strategic method, showing, for various budgets, the difference in risk reduction achieved by the strategic method compared to a number of benchmark strategies.

The strategic method computes a strategy which is optimal if the various inputs to the strategic method are correct. Since it is very hard to know in practice exactly what the probability of failure or the consequent loss costs either before or after risk mitigation, the strategy recommended by the strategic method may turn out to be less than optimal. More importantly, the strategy recommended by the strategic method could turn out to be worse than other fixed reasonable strategies less sensitive to the accuracy of the input estimates. The basic question is, "Will the strategic method perform worse than other strategies if given inaccurate inputs?"

In order to examine this question, we have carried out a number of experiments in which uncertainty/noise is introduced into the inputs to the strategic method, and then the risk reduction achieved by the recommended strategy compared with the risk reduction achieved by a fixed strategy. The benchmark strategies we use are:

1. A random (or arbitrary choices of attribute and techniques and the order in which they are applied) strategy.

2. The cheapest strategy.

3. In subsequent experiments, a 'great' strategy.

Figure 1, reproduced from [Port et al 2005] compares the risk reduction achieved using various different strategies. Risk reduction and cost are assessed according to the data in Tables 1-3.

As expected theoretically, the RE profile for the random strategy is approximately linear and not very appealing. We can assess the effectiveness of a strategy by considering how well it reduces risk for a given cost as compared to a random strategy. Also as expected, the figure clearly shows that the strategic method strategy reduces overall RE as a function of cost better than all other strategies.

3.1 Calculating the difference between strategies

Recall that a strategy is a sequence of pairs $<a,t>$ recommending application of technique t to attribute a. For each technique-attribute pair, we calculate risk reduction using the following values:

$P_{before}(a)$, the probability of loss for attribute a before application of any risk reduction.

$S_{before}(a)$, the cost consequent to the loss for attribute a.

$C(a,t)$, the cost of applying technique t to attribute a.

$P_{after}(a,t)$, the probability of loss for attribute a after application of technique t.

$S_{after}(a)$, the cost consequent to the loss for attribute a.

The risk reduction achievable from a strategy $<a_1,t_1>$, $<a_2,t_2>$, ..., $<a_n,t_n>$ using some budget b is the risk reduction achieved by applying the techniques t_i to the attributes a_i in the order specified in the strategy up until the point where the budget has been exhausted.

More precisely, we define the risk reduction achieved by applying the first k techniques of a strategy $S=<a_1,t_1>$, $<a_2,t_2>$, ..., $<a_n,t_n>$ to be $\delta RE^k(S) = \sum_{i=1..k} P_{before}(a_i) * S_{before}(a_i) - P_{after}(a_i,t_i) * S_{after}(a_i)$ and the associated cost $C^k(S) = \sum_{i=1..k} C(a_i,t_i)$. The risk reduction achieved at a budget b, denoted $\delta RE(b,S)$, is value of $\delta RE^k(S)$ for the largest k such that $C^k(S) \leq b$. In this paper, we are only interested in *risk reduction*, not in the total absolute risk exposure.

In order to evaluate the effect of uncertainty on the inputs to the strategic method, we will conduct experiments in which we generate an optimal strategy using the strategic method using some known values for the input to the strategic method, but evaluate the risk reduction achieved by that strategy assuming values of the risk probability and cost matrices which are perturbed from those input values.

The quantity we perturb in our experiments is the *effectiveness* of technique t_i on attribute a_j, defined to be $\rho_{ij}=(P_{before}-P_{after})/P_{before}$. If $\rho_{ij}=0$, then the technique in

completely ineffective. If $\rho_{ij}=1$, then the technique completely mitigates the risk for attribute a_j. If $S_{before}=S_{after}$ (as it is in all our experiments here), then the risk reduction achieved from $<t_i,a_j>$ is $\rho_{ij}*P_{before}(a)*S_{before}(a)$. Let our estimate of the effectiveness matrix be ρ^0. We perturb the effectiveness matrix by multiplying it by Gaussian noise, replacing each entry ρ^0_{ij} with $\rho^0_{ij}*N(1,\sigma^2)$, (If this would make $\rho_{ij}<0$, we replace it with 0, if it would make $\rho_{ij}>1$, we replace it with 1).

The question of how to perturb a number representing an effectiveness in the range [0,1] is an important one, and there could be a concern that the way we have truncated the normal distribution could bias the results of our experiments. We repeated some of the experiments in Section 3. using different mechanisms for perturbing effectiveness. We tried two schemes: 1) $\rho_{ij} = \rho^0_{ij}+N(0,\sigma^2)$, with no cutoff, with the result that some of the perturbed effectiveness numbers may be less than 0 (a risk reduction technique is worse than useless), or greater than 1 (a risk reduction technique beyond our wildest dreams), and 2) using an accept-reject method, setting $\rho_{ij} = \rho^0_{ij}+N(0,\sigma^2)$, but rejecting and resampling if this puts ρ_{ij} outside the range [0,1]. Once accepted, ρ_{ij} is guaranteed to be in the range [0,1]. The first scheme did not seem to affect the mean but did appear increase the variance of the experiment results. The second scheme did not appear to have an appreciable effect. Further work can be carried out in this area, but we believe that the mechanism we have chosen for adding noise in this paper does not appreciably bias the results.

3.2 Strategic method versus random strategy

We generate a random strategy S by applying techniques from $T=\{T_1,...,T_m\}$ to attributes from $A=\{A_1, ..., A_n\}$ as follows:

1. Initially set *S* to be the empty sequence.
2. Randomly select an attribute *a* from *A*. Delete *a* from *A*.
3. Randomly select a technique *t* from *T* such that *t* is applicable to *a*, i.e. $P_{after}(a,t)\neq X$ and $C(a,t)\neq X$.
4. Add $<a,t>$ to the end of *S*.
5. Repeat from (2) above until *A* is empty.

Using the above procedure, we generate a random strategy once per experiment, and call it S_{fixed}.

We now calculate the difference in risk reduction achieved by S_{fixed} compared to the strategy recommended by the strategic method, S_{opt}, for various values of budget and various amounts of noise in the effectiveness matrix.

We compare the random strategy to the strategic method strategy as follows:

1. Pick fixed budget $b \in \{50, 100, 150, 200, 250, 300, 350, 400\}$.
2. Pick noise level $\sigma \in \{0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.7, 1.0\}$
3. N (=1000) times:
4. Add noise to effectiveness matrix: $\forall i,j$, $\rho_{ij}=||\rho^0_{ij}+N(0, \sigma^2)||$
5. Evaluate %age difference between S_{opt} and S_{fixed}, $\Delta=(\delta RE(b, S_{fixed})- \delta RE(b, S_{opt}))/ \delta RE(b, S_{opt})$

Add a point to the plot with *x* coordinate σ, *y* coordinate the mean value of Δ, and if desired add error bars to that point to indicate the standard deviation in Δ.

Figure 2 below shows the results of comparing the optimal with random strategies from two typical experiment runs. The quantity on the Y axis is the difference between the risk reduction achieved by S_{fixed} and S_{opt} expressed as a percentage of the risk reduction achieved by S_{opt}. For example, if S_{opt} reduces risk exposure by 130, and S_{fixed} reduces it by 75, then we plot (75-130)/130 = -42%. Negative numbers on the Y axis indicate that the optimal strategy reduces risk exposure more than (i.e. is better than) the random strategy. From our experiment we note the following:

The optimal strategy is *always* much better than the random strategy regardless of the noise and budget.

As the amount of noise in the effectiveness matrix increases, the difference between the optimal and random strategies decreases, but the standard deviation increases, i.e. the optimal strategy behaves more and more like a random strategy.

Higher budgets result in less difference between the optimal and random strategies, i.e. as we spend more money, the random strategy becomes more optimal. That is, spending more money is always a way to mitigate risk.

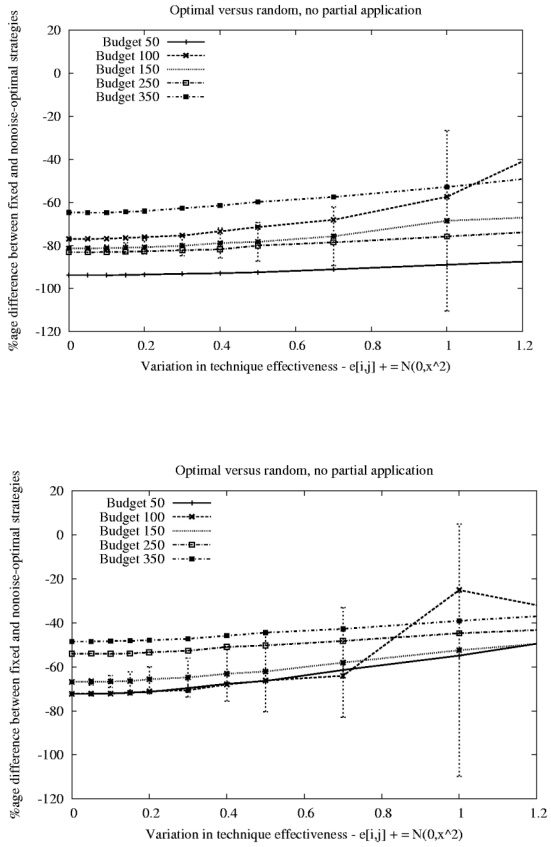

Figure 2: results from a typical experiment run comparing optimal with a random strategy for various budgets and amounts of noise in effectiveness matrix. Error bars indicate standard deviation for the budget=100 curve.

3.3 Strategic method versus the cheapest strategy

The strategic method strategy is typically much better than random strategies, but the random strategies are generally so poor that many strategies are not surprisingly better. A more interesting benchmark strategy to use is the 'cheapest' strategy, which selects for each attribute the cheapest technique applicable to that attribute, and applies those attribute-technique pairs in order of increasing cost. We generate the cheapest strategy applying techniques from $T=\{T_1,...,T_m\}$ to attributes from $A=\{A_1, ..., A_n\}$ as follows:

1. Initially set S to be the empty sequence.

2. Select the first attribute a from A. Delete a from A.

3. Select the cheapest technique t^* from T such that t^* that is applicable to a, i.e. $P_{after}(a,t^*) \neq X$ and $C(a,t^*) \neq X$, and $C(a,t^*) \leq C(a,t)$ for all techniques t.

4. Add $<a,t>$ to the end of S.

5. Repeat from (2) above until A is empty.

Sort the technique-attribute pairs from S in order of increasing cost.

Figure 3 below shows the results of benchmarking the optimal against the cheapest strategy. Each point in the lower graph represents the difference between the optimal and cheapest strategy for a certain budget and noise level, averaged over 1000 runs. The upper graph in Figure 3 shows the results of individual runs rather than averages. There are 100 points for each noise level, and the noise level ranges from 0.0 to 1.0 as before. In this experiment we note the following:

1. Even for high levels of inaccuracy in the effectiveness matrix, the strategic method strategy is *significantly* better than the cheapest strategy.

2. Again, the performance of the optimal strategy degrades as the amount of inaccuracy in the effectiveness matrix increases.

3. Interestingly, from budget 100 onwards, as we increase the budget, the performance of the cheapest strategy gets *worse* compared to the optimal strategy, not better as was the case in Figure 1.

For low budgets, the standard deviations become extremely large (much larger than in Figure 1) when there is a lot of inaccuracy in the effectiveness matrix – when we are adding noise from N(0,1) to the effectiveness matrix, the standard deviation of the %age difference between the risk reductions achieved by the optimal and cheapest strategies are 113, 97, 300, 39, 30 at budget 50, 100, 150, 250, 350 respectively. This increase in standard deviation can be seen clearly in the rightmost graph of Figure 2. Based on these results in the presence of high noise, we would not expect the optimal strategy to necessarily perform better than the cheapest strategy for low budgets, even though on average it may perform significantly better for low budgets.

3.4 Strategic method versus the Great Strategy

A final challenge, which also demonstrates another interesting point, is to compare the optimal strategy against a strategy which is not just good, but great. We can choose such a great strategy by applying the strategic method itself.

In this final set of experiments, we first apply the strategic method to generate an optimal strategy – in fact, the optimal strategy shown in Table 4, which is our 'great' fixed strategy S_{fixed}. We then perturb the loss potentials of Table 1and calculate a new optimal strategy, S_{opt}. Comparing S_{opt} and S_{fixed} in this way provides us of a measure of how well the strategic method adjusts the strategy to changes in loss potential.

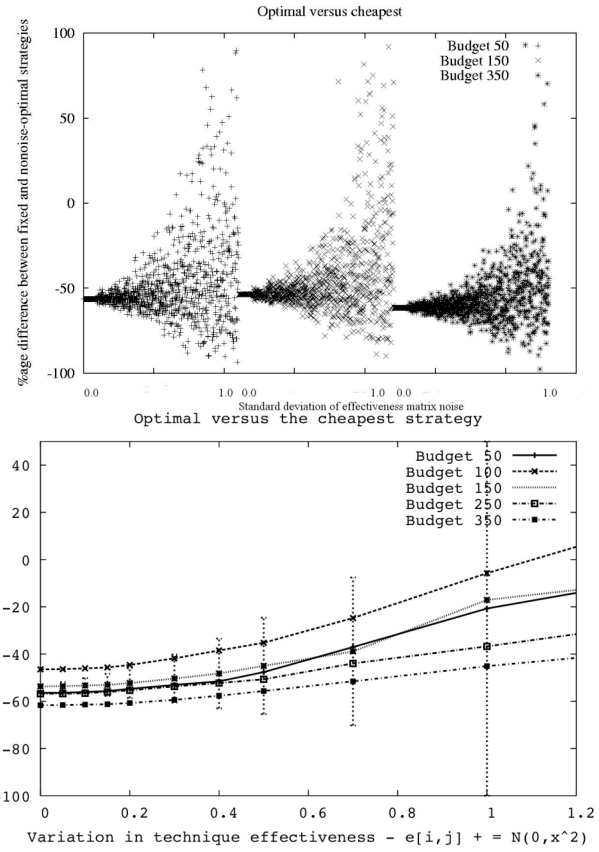

Figure 3: optimal versus cheapest strategy – points in the graph at the top are single samples, points in the graph below are single samples.

We perturb each potential independently of the others, multiplying each loss potential by 0 (31% of the time), 1 (38% of the time) and 1.5 (31% of the time). These possibilities correspond to scenarios where the loss potential for an attribute differs from the norm either because (0) an attribute risk is inapplicable or has already been completely mitigated, (1) the loss potential is normal, or (2) the loss potential is elevated compared to normal.

Figure 4 below shows the results of benchmarking the 'great' strategy against the strategic method strategy chosen for a random perturbation of the loss potentials. We note the following:

The optimal strategy is significantly better than the 'great' strategy. With no noise, the improvement is in the 0-15% range.

As the amount of noise increases, the performance of the optimal strategy improves compared to the 'great' strategy.

The larger the budget, the better the optimal strategy performs over the 'great' strategy.

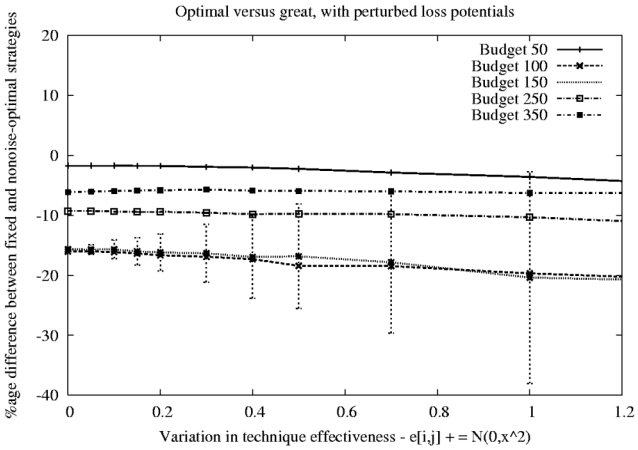

Figure 4: 'great' strategy against strategic method strategy for random perturbation of loss potentials.

4. Experiments with an Alternative Optimization Method

In parallel with development of the strategic method [Kazman, Port et al], a conceptually similar approach has been pursued in the form of the Defect Detection and Prevention (DDP) process [Feather&Cornford, 2003]. DDP has been used to help assess risks and plan their cost-effective mitigation primarily for space mission technologies [Feather et al, 2005]. This section outlines DDP and its similarities to and differences from the strategic method , and describes some experimental evaluations of DDP-selected risk reductions.

DDP, like the strategic method , treats risks as probabilistic events whose occurrence would detract from the attainment of attributes of interest, and treats techniques (with associated costs) as options to consider for reducing the likelihoods and/or severities (degree of impact) of risks. DDP can be used much like the strategic method to determine an "optimal" risk-reducing strategy – that is, for a given budget say, determine the selection of techniques that minimize overall risk while remaining within budget. Differences between the strategic method and DDP lie in some of the internal calculations of risk reduction, and the means by which each technique arrives at an "optimal" risk-reducing strategy.

DDP allows for the possibility that applying a technique may reduce several risks, and that a risk may be reduced by several of the techniques being applied. While the strategic method assumes one technique per risk, this can be accounted for by splitting the risk into sub-risks and adding dependencies. In DDP the effect of a technique at reducing a risk as a number in the range 0 – 1, the proportion by

which the risk will be reduced. DDP's rule for computing the combined effect of two techniques at reducing the same risk is to treat them as "filters" operating in "series": e.g., if Technique A reduces a risk by 0.4 and Technique B reduces that same risk by 0.3, then applying both of them works as follows: suppose the initial risk value (likelihood x severity) = R; Technique A reduces the risk by 0.4, yielding a reduced risk value of (1 – 0.4) *R = 0.6R; Technique B then reduces that by 0.3, yielding a reduced risk value of (1 – 0.3) * 0.6R = 0.42R. Note that the order of the techniques doesn't matter; the other way around would change the risk R to 0.7R to 0.42R, the same end result.

A fundamental difference between DDP and the strategic method is that DDP does not output a strategy. Rather it determines a set of attribute-technique pairs that optimize risk-reduction with respect to fixed budget. The order in which the attribute-technique pairs are executed is not considered. Fortunately for our sensitivity studies at present, order is not relevant.

DDP treats determining an optimal risk-reduction selection as a classic optimization problem – for a given selection of techniques, DDP can be used to compute the cost (the sum total cost of that selection), and benefit (the sum of the attributes as reduced by the risks that detract from them, where those risks have been reduced by the selected techniques). By default, DDP treats the techniques as independent options, which means for N techniques, there are 2N ways of selecting from among them. DDP then uses either exhaustive search when the number of techniques is small (16 or so, depending on user patience!), or heuristic search to locate near-optimal solutions (the current DDP implementation uses simulated annealing for this purpose).

In addition to the evaluations of the strategic method described in the other sections of this paper, the first author (Richardson) had done some similar evaluations of a hybrid of strategic method and DDP using a slightly smaller dataset comprising 12 risks and 7 techniques (in the interests of remaining within the space limits, we do not reproduce this dataset herein). This inspired the last author (Feather), the primary developer of DDP, to recreate those experiments as closely as possible in DDP. For the dataset in question, Richardson had used the approach of the strategic method to compute a "Great" strategy – an ordering of the techniques, to be used in the following way: for the given budget, consider each technique in the ordering; if enough budget remains to pay for that technique, select it, decrease the remaining budget by that technique's cost, and move to the next technique in the order; conversely, if insufficient budget remains to pay for that technique, then assume that the technique can be partially applied, in the amount of the fraction computed by dividing the budget remaining by the cost of the technique (and since this will use up all the remaining budget, there is no need to continue considering techniques). For example, if the technique in question is static analysis of code, it is reasonable to expect that applying static analysis to a portion of the code will reveal a portion of the errors detectable by static analysis. Thus for a given budget, Richardson's approach yields some number of techniques to be applied in full, and possibly one to be applied fractionally (in some cases the last considered technique uses up the entire remaining budget, in which case no fractional applications result). Feather modified DDP's exhaustive search for optimal risk reduction to also encompass partial (fractional) application, and recreated as closely as possible Richardson's dataset within DDP. It was revealed that Richardson's strategic method derived "Great" strategy, and DDP's exhaustive search, always (for the budget levels considered in the experiments) computed the same selection of techniques. Feather then performed some preliminary evaluation experiments akin to those described in section 3.4 to compare the "Great" strategy, predetermined ahead of time, against an optimal (as located by DDP) strategy computed after perturbing loss potentials for the 12 risks. The same perturbations as described earlier were used: each risk is perturbed potential independently of the others, multiplying its loss potential by 0 (31% of the time), 1 (38% of the time) and 1.5 (31% of the time). Having determined the two strategies, gaussian noise was used to perturb the effectiveness matrix in each of the experimental runs, to mimic uncertainty in effectiveness numbers (again, as described earlier in the paper). The results from one set of these experiments are see in figure 4

The same approach as used elsewhere in this paper is used to plot the mean percentage difference between the approaches being compared – here, between the predetermined and post-risk-perturbation-determined strategies. This is done for a several budgets (the different colors of points) and several noise levels (the horizontal axis). The plot reveals that the post-risk-perturbation-strategy is, on average, superior to the predetermined one. This is as we would hope: it indicates that taking into account additional information about relative prevalence of risks indeed improves our ability to select a better strategy. We also see the phenomenon of increasing budgets diminishing the differences – as more and more risk reduction techniques can simultaneously be afforded, selection from among them becomes less critical. Finally, it seems that in this case, increasing the noise level somewhat diminishes the difference between the two strategies.

A separate experiment that disallowed use of partial "fractional" techniques showed a greater leaning in favor of DDP, perhaps because the budget thresholds (at intervals of 10) had an unfortunate interaction with one of the techniques (costing 5), leading to "wastage" of budget when naively following the strategic method of determination.

These DDP experiments are just preliminary, so we do not yet draw any strong conclusions from them. The do, however, indicate that the evaluation approaches described in this paper can yield insights when applied to alternate risk-reduction models, such as "standard" DDP, and DDP

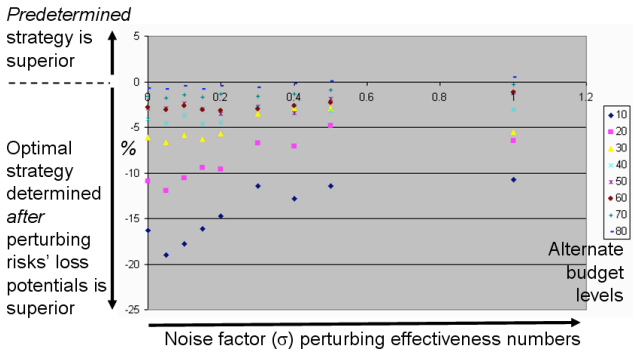

Figure 4: 'great' strategy against DDP strategy for random perturbation of loss potentials.

augmented with Richardson's concept of partial application of techniques to consume remaining budget.

5. Conclusions

The strategic method calculates risk reduction strategies specific for a project, tailored to the amounts of risk and possible risk reduction in that project. In fact \cite{Port et al, 2005}, the computed strategies are optimal as long as the inputs to the strategic method do accurately reflect the amounts of risk and possible risk reduction in that project.

In practice, accurately estimating quantities such as probability of failure, cost consequence of failure, and effectiveness of risk mitigations is very hard.

In this paper, we examined the effect that inaccuracies in the inputs to the strategic method have on the risk reduction achieved by the strategic method strategy compared to various baseline strategies.

The strategic method strategy generally performs much better than random strategies, and that improvement can increase as our assessment of the effectiveness of available mitigations becomes more inaccurate. We speculate that the reason for the latter is that the strategic method strategy is in some sense 'robust' in the presence of noise, since it tackles the most important risk attributes early, even if the amount of mitigation achieved by a mitigation technique may differ significantly from our expectations.

The strategic method strategy performs much better than the cheapest strategy, although the improvement in performance decreases as our assessment of the effectiveness of available mitigations becomes more inaccurate. For some values of budget and injected noise, the standard deviation of the measured improvement is extremely large. For those values, we would not necessarily expect the optimal strategy to perform better than the cheapest strategy, even though it will do so on average.

When we randomly perturb the loss potentials input to the strategic method, the strategic method strategy – which is tailored to those values of loss potential – performs better than a fixed 'great' strategy (the strategy of Table 4). The size of the improvement increases as our assessment of the effectiveness of available mitigations becomes more inaccurate.

Using the strategic method to compute optimal strategies does seem to result in worthwhile – sometimes very large – improvements in risk reduction, even when our assessment of the effectiveness of available mitigations is inaccurate.

Our results have been cross-validated by considering the alternative risk-reduction versus cost optimization method DDP.

6. Acknowledgements

This research was carried out at the NASA Ames Research Center, the University of Hawaii, and the Jet Propulsion Laboratory, California Institute of Technology, partly under a contract with the National Aeronautics and Space Administration and partly supported by JAXA. This work was partially funded by NASA Cooperative Agreement NNA05CS31A and partially by NASA's Exploration Systems Mission Directorate.

7. References

[Feather&Cornford, 2003] Feather, M.S., and Cornford, S.L., "Quantitative risk-based requirements reasoning", Requirements Engineering (Springer), 8(4), pp 248-265, 2003.

[Feather et al, 2005] Feather, M.S., Cornford, S.L., Hicks, K.A. and Johnson, K.R., "Applications of tool support for risk-informed requirements reasoning", Computer Systems Science and Engineering (CRL Publishing Ltd); 20(1): 5-17, Jan 2005.

[Port et al] Dan Port, Rick Kazman, Blanca Polo, Haruka Nakao, Masa Katahira, "Practicing What is Preached: 80-20 Rules for Strategic IV&V Assessment", technical report CSSE-TR20051025, Center for Strategic Software Engineering, University of Hawaii at Manoa, 2005.

8. BIOGRAPHIES

Julian D. C. Richardson is a Research Scientist working for RIACS in the Reliable Software Engineering Group at NASA Ames Research Center. He was awarded a PhD in the field of Artificial Intelligence from the University of Edinburgh, Scotland, in 1995. His main research interests are software risk assessment, verification and validation, and automated software engineering. He is a member of the ACM.

Martin S. Feather is a Principal in the Software Quality Assurance group at JPL. He works on developing research ideas and maturing them into practice, with particular interests in the areas of early phase requirements engineering and risk management and of software validation (analysis, test automation, V&V techniques). He obtained his BA and MA degrees in mathematics and computer science from Cambridge University, England, and his PhD degree in artificial intelligence from the University of Edinburgh, Scotland. For further details, see http://eis.jpl.nasa.gov/~mfeather

Dan Port is a Professor of Information Technology Management at the University of Hawaii's Shidler College of Business and a Visiting Associate at the Center for Software and Systems Engineering at the University of Southern California. His research focuses on strategic planning and assessment of IV&V activities, strategic software engineering, empirical software engineering, and software engineering education. Daniel attended UCLA and graduated with a degree in Mathematics and later received his Ph.D. in Applied Mathematics and Theoretical Computer Science from Massachusetts Institute of Technology.

9. APPENDIX

The attribute and technique descriptions for the data indicated in the tables are:

A1: Robustness/Independent redundancy (No Single Failure Point, Priority Inversion)
A2: Robustness/Independent redundancy (No Single Failure Point, Requirement Consistency, Completeness)
A3: Robustness/Independent redundancy (No Single Failure Point, Code Quality)
A4: Stability of Performance (Timing, Message queue over flow)
A5: Real time performance (Don't skip the data flame)
A6: Development Schedule
A7: Cost
A8: Portability/ Replaceability, Adaptability (to Hardware or Driver)
A9: Maintainability/Changeability
A10: Scalability (Capability of adding application code)
A11: Testability
A12: Understandability (access to code)
A13: Resource Utilization (How much resource used when maximum process is on using past system
A14: Vender Support (Response time)

T1: Test Suites	T8: Custom Method
T2: Analysis Using Model	T9: Interview Vendor
T3: API Test	T10: Investigation of past data
T4: Model Checking	T11: Test on Emulator
T5: Code Review Lessons Learned	T12: Best Guess
T6: Static Analysis of code	T13: Benchmark test
T7: Estimation	T14: Simulation